NOVEMBER/DECEMBER 2025

23PCS34 — DATA SCIENCE AND ANALYTICS

Time : Three hours Maximum : 75 marks

Shanmuga Industries Arts & Science College T.V.Malai

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

1. Define the term Ecosystem.
2. List any two machine learning algorithms.
3. What is Data analytic tool?
4. Define the term Data review process.
5. List and define some basic Data types in R.
6. What is Data exploration?
7. Define the term K-means clustering.
8. What is Decision tree algorithm in R?

9. List some real-life applications of Artificial Intelligence.

10. What is meant by Clustering in AI?

SECTION B — (5 × 5 = 25 marks)

Answer ALL questions.

11. (a) Illustrate the several facets of data in Data science.

Or

(b) Explain the steps in Data modeling of machine learning.

12. (a) Describe the types of Data analytics.

Or

(b) Write the benefits of advanced Data analytics.

13. (a) How to import data in R? Discuss with example.

Or

(b) Elucidate Exploratory data analysis using R.

Shanmuga Industries Arts & Science College T.V.Malai

14. (a) Illustrate the types of clustering in R programming.

Or

(b) Why is "smoothing" usefu1 when applying Naïve bayes? Explain.

15. (a) Differentiate machine learning and deep learning.

Or

(b) Write about association rules, fundamentals.

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. What is Bigdata? How is it different from Data science and Data analysis? Discuss.

17. Explain the data analytics life cycle phases.

18. Describe the process of Descriptive statistics in R.

19. Explain, how to use Naive Bayes classifier in R in detail.

20. Discuss with example about Logistic regression.

———————